International Journal of Agricultural Science and Research (IJASR)
ISSN(P): 2250-0057; ISSN(E): 2321-0087
Vol. 8, Issue 1, Feb 2018, 129-136

ARIMA FORECAST OF AREA, PRODUCTION AND PRODUCTIVITY OF RICE IN AMALGAMATED RAIPUR USING R

G. BHAVANI[1] & A. K. SINGH[2]

[1]*Student M.Sc. (Agri), Department of Agricultural Statistics, IGKV, Raipur, Chhattisgarh, India*

[2]*Professor and Head, Department of Agricultural Statistics, IGKV, Raipur, Chhattisgarh, India*

ABSTRACT

The forecast of the production of rice over future time series in Chhattisgarh has always been a matter of concern to the planners of the State, especially being a predominantly tribal region. The present study aims at patterns of studying area, production and productivity of rice crop for Amalgamated Raipur districts of Chhattisgarh Plains over a time series from 1980 to 2015 and their forecasts for the next five years using ARIMA and double exponential methods. The forecasts undertaken were based on the libraries developed in the globally famous statistical software environment "R: The R Project for Statistical Computing" available at www.r-project.org. Some good models have been developed for the area, production and productivity wherever model diagnostics permitted. A synopsis and evaluation of Good fitted forecast models have been summarized. Subsequently, forecast based on these good models have been attempted from 2015-16 to 2019-20.

KEYWORDS: Amalgamated Raipur, R, AR, MA, ARMA, ARIMA, ARIMA (p,d,q), DASC, Forecasting, CI80%, CI95%

Received: Jul 02, 2017; **Accepted:** Jul 23, 2017; **Published:** Feb 02, 2018; **Paper Id.:** IJASRFEB201818

Original Article

INTRODUCTION

Chhattisgarh State was created on November 1, 2000 by incorporating 16 districts of undivided eastern Madhya Pradesh. It shares its borders with six states of the country - Orissa in the east, Jharkhand in the north-east, Madhya Pradesh in the west and north-west, Uttar Pradesh in the north, Maharashtra in the south-west and Andhra Pradesh in the south and south-east. With a geographical area of 137.90 lakh ha (www.agridept.cg.gov.in), Chhattisgarh is the tenth largest State (https://en.wikipedia.org) in India, contributing about 2.97 per cent of the total annual food grains production in India (DASC, 2012-13). The State has a net sown area 47.75 lakh ha, cropping intensity 138 percent and forest coverage of 63.53 lakh ha. After creation of Chhattisgarh, many districts, and many erstwhile districts were divided, among which Raipur was divided twice in 1998-99 and 2011-12. Due to these divisions the divided parts of the corresponding districts suffered due to data availability in order to determine the patterns of area, production and productivity of different crops. So, to get a clear picture, the areas and productions of the divided parts were amalgamated, and productivity averaged, to form Amalgamated Raipur, as if they were not divided. Thus, the entire study and forecast were made for this Amalgamated Raipur.

METHODS

The present study utilizes a relational database created in the Department of Agricultural Statistics and Social Science, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh, on the basis of data acquired from www.agridept.cg.gov.in and Anonymous (1981-1998). The detailed methodologies adopted in the present

investigation are presented in the following sub-heads.

ARIMA MODEL

In time series analyses, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model, Anonymous (2016). These models are fitted to time series data either to better understand the data or to predict future points in the time series (forecasting). They are applied in some cases where data show evidence of non-stationary and where an initial differencing step (corresponding to the "integrated" part of the model) can be applied to reduce the non-stationarity. Thus, the autoregressive (AR) component of the ARIMA model is a representation of a type of random process representing certain time-varying processes, economical phenomenon of the nature, etc. It specifies that the output variable depending linearly on its own previous values and on a stochastic term. This AR(p), an autoregressive model of order p, can be defined as,

$$Y_t = \mu + \sum_{i=1}^{p} \Phi_i Y_{t-i} + \mathcal{E}_t \tag{1}$$

where $\Phi_1, \ldots . \Phi_p$ are the parameters of the model, μ is a constant, and $\mathcal{E}_t$ is stochastic white noise.

The second component of ARIMA is Moving-Average (MA) model. The general MA process of order q can be defined as,

$$Y_t = \mu + \mathcal{E}_t + \Theta_1 \mathcal{E}_{t-1} + \cdots + \Theta_q \mathcal{E}_{t-q} \text{-} \tag{2}$$

where μ is the mean of series, the $\Theta_1, \ldots, \Theta_q$ are the parameters of the model and the $\mathcal{E}_t, \mathcal{E}_{t-1}, \ldots, \mathcal{E}_{t-q}$ are the stochastic noise error terms. The value of q is called the order of the MA model.

Now, combining the AR process with the MA process and integrating, with differencing step, to remove non-stationarity from the non-seasonal time series, the ARIMA is generally denoted by ARIMA (p,d,q), where parameters p, d, and q are non-negative integers, p being the order of the autoregressive model, d being the degree of differencing, and q being the order of the moving-average model. In other words, ARIMA (p,d,q), is an autoregressive integrated moving average time series.

Thus, the ARIMA (p,d,q) model, (Qbal, N. et al. 2005) can be represented by the following general forecasting equation:

$$=\mu + Y_{t-i} \qquad +_{t-j} \qquad + \tag{3}$$

Or,

$$=\mu + Y_{t-i} \qquad +_{t-j} \tag{4}$$

DOUBLE EXPONENTIAL SMOOTHING MODEL

The double exponential smoothing is best applied to time series that exhibit prevalent additive (non-exponential) trend, but do not exhibit seasonality. (Prajakta S. (2004)). The recursive form of the Holt's double exponential smoothing

equation is expressed as follows:

$$\widehat{F_t}(m) = S_t + m \times b_t$$

$$S_t = \alpha Y_t + (1-\alpha)(S_{t-1} + b_{t-1})$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_t$$

Where,

is the value of the time series at time t;

is a smoothed estimate of the value of the time series Y at the end of period t;

is a smoothed estimate of average growth rate at the end of period t.

α is the level smoothing coefficient.

β is the trend smoothing coefficient.

is the m-step-ahead forecast values for Y from time t.

The estimates and forecasts for above mentioned models have been computed based on the libraries developed in the globally famous statistical software environment "R: The R Project for Statistical Computing" available at www.r-project.org. The authors have used the methods relevant to these as illustrated in Coghlan, A. (2014, 2017).

RESULTS AND DISCUSSIONS

Based on the methods elucidated above, good models have been estimated using double exponential and ARIMA models as listed in following Table 1.1, for which predictions were made:

Table 1.1: Synopsis and Evaluation of fitted models based on (i) Model-1: Double exponential model with parameters α, β, a, b and (ii) Model-2: ARIMA model with parameters ACF, P-value, AIC, BIC, σ^2; and respective diagnostic plots with confidence intervals.

Crop	Variables	Region/District	Model Adequacy
Rice	Area	Amalgamated Raipur	Model-1:Good fit
			Model-2:Good fit
	Production	Amalgamated Raipur	Model-1:Good fit
			Model-2:Good fit
	Productivity	Amalgamated Raipur	Model-1:Good fit
			Model-2:Good fit

So, using the good models as indicated in Table 1.1 above, the estimated parameters of forecasting models for area, production and productivity under the rice crop for Amalgamated-Raipur for the period 1980-81 to 2014-15 have been presented along with the relevant forecasts and their confidence intervals in Figure .1 to Figure .6 in the Appendix-A, respectively, for double exponential smoothing forecasting as well as for ARIMA forecast. These are followed by respective graphs for plots of model diagnostics of (a) observed *vs*. fitted and (b) double exponential smoothing forecast and (c) Time Series plot and (d) ARIMA Forecasting etc.

The figures Figure .1 from Figures.6 in the Appendix-A, can be perused for various estimates of parameters for levels, trends and seasons of area, production and productivity of Amalgamated-Raipur under rice crop. Thus, from

Figures..1 in the Appendix-A, left side table, for area of rice it is found that the α coefficient of double exponential model is very high (1, equal to maximum 1), therefore the level of this model can be said to depend more on the recent observations than the past observations during the period of study from 1980-81 to 2014-15, with the start of the level at 772.3, whereas the value of β-coefficient is very low (0.05, close to zero) indicating that there is almost no trend in the double exponential model estimate, with the start of the trend at a low value of 3.96, for the period 1980-81 to 2014-15. Secondly, from Figures.2, left side graph, it is clear that the autocorrelation (i.e. ACF) is zero because p-value > 0.80, from Figure.1.1, left side table, the AIC/BIC values are close to 371-373 with variance of 2254, i.e., standard deviation of 47.47 thousand ha.. The forecast of the area under rice by double exponential model and ARIMA model are depicted in Figures.1 and Figures.2 in the Appendix-A, respectively, for 80% and 95% confidence intervals for five years from 2015-16 to 2019-2020. On perusing these graphs it may be concluded that forecasted area for the double exponential model are in the range of (776.3069 to 792.1785) 000'ha with approximate 80% confidence interval (CI80%) of (630-930) 000'ha and approximate 95% confidence interval (CI95%) of (550-1050) 000'ha. On the other hand, the forecasted area for ARIMA model is 772.339 000'ha for all the five years from 2015-16 to 2019-2020 with CI80% values being in the range 630-900 000'ha and CI95% range 550-1000 000'ha. Clearly, the ARIMA estimates give lower confidence intervals with pointed better pointed forecast.

From Figures.3, left side table, for production of rice it is found that the α coefficient of double exponential model is low (0.48, equal to maximum 1), therefore the level of this model can be said to depend more on the both recent observations, with the start of the level at 1545.4, whereas the value of β-coefficient is very high (0.975) indicating that there is trend in the double exponential model estimate, with the start of the trend at a value of 179.9, for the period 1980-81 to 2014-15. Secondly, from Figures. A.4 in the Appendix-A, left side graph, it is clear that the autocorrelation (i.e. ACF) is zero because p-value > 0.80, from Figures.1.3, left side table, the AIC/BIC values are close to 472.71- 491.9 with variance of 33149, i.e., standard deviation of 182.06 000' tonnes. The forecast of production under rice by double exponential model and ARIMA model are depicted in Figures.3 and Figures.4 in the Appendix-A, respectively for 80% and 95% confidence intervals for the five years from 2015-16 to 2019-2020. On perusing these graphs it may be concluded that forecasted production for the double exponential model are in the range (1725.416 to 2445.13) 000'tonnes with approximate 80% confidence interval (CI80%) of (1200-3600) 000'tonnes and approximate 95% confidence interval (CI95%) of (600-4200) 000'tonnes. On the other hand, the forecasted production for ARIMA model are in the range 1454.513- 1531.511 000'tonnes for all the five years from 2015-16 to 2019-2020 with CI80% values being in the range 1000-2000 000'tonnes and CI95% range 700-2300 000'tonnes. Evidently, the ARIMA estimates give better pointed forecast with lower confidence intervals.

Next, from Figures.5 in the Appendix-A, left side table, for productivity of rice, the α coefficient of double exponential model is low (0.51, equal to maximum 1), so the level of the model depends more on the recent observations, starting at 1545.4, the β-coefficient being very low (0.43) indicating no trend in the double exponential model estimate, with the intercept of 1956.7, for the period 1980-81 to 2014-15. Secondly, from Figures.6 in the Appendix-A, left side graph, the autocorrelation (i.e. ACF) is zero because p-value > 0.80, from Figures.5, left side table, the AIC/BIC being close to 490.08 – 493.19 with variance of 63355, i.e., standard deviation of 251.70 kg. The forecast of productivity under rice by double exponential model and ARIMA model are depicted in Figures.5 and Figures.6 in the Appendix-A, respectively for 80% and 95% confidence intervals for the five years from 2015-16 to 2019-2020. On perusing these graphs it is concluded that productivity for the double exponential model are forecasted in the range of 2036.08 to 2353.25

kg with approximate CI80% of 1200-3400 kg and CI95% of 700-3900 kg. On the other hand, the forecasted productivity for ARIMA model are in the range 1858.49 - 1934.559 kg for all the five years from 2015-16 to 2019-2020 with CI80% values being in the range 1400-2400 kg and CI95% range 1100-2700 kg. On comparison, the ARIMA estimates gives lower confidence intervals with better pointed forecast.

Finally, based on Figures .1-.6, a clear picture emerges in terms of a great diversification of the area under rice to other crops from the year 1998-99 onwards in Amalgamated Raipur, which by the way coincides with the first split of Raipur district in the same year 1998-99 too. Due to these factors the area and production of rice take a sudden dip in 1998-99 and then gradually rise onwards. However, the productivity maintained a rising trend after a few years of stationary following diversification of rice area to other crops in Raipur district and the first split of Raipur district, in 1998-99, possibly as a result of policy paralysis emerging out of shock due to above-said diversification along with a split in the district. Despite these, the production of rice in Amalgamated Raipur in 2014-15 has now crossed the peak of rice production compared to what it was just before the start of diversification, in 1998-99, of rice area to other crops in Raipur district.

CONCLUSIONS

It may be concluded from the present study that ARIMA models fitted better than the double exponential models uniformly, as depicted from Figures .1-.6 in the Appendix-A, for area, production and productivity of rice under Amalgamated Raipur, in terms of all the standard diagnostics' estimates like AIC, BIC, p-values, confidence intervals (CIs), autocorrelation functions (ACFs) on the one hand; and graphical diagnostics on the other hand, like observed vs fitted graphs and ARIMA ACF graphs. It is further concluded that the area and production of rice take a sudden dip in 1998-99 indicating a great deal of diversification of rice area to other crops. However, both parameters gradually rise onwards. Yet, the productivity maintained a rising trend, after a few years of stationary following diversification in the rice area to other crops in Raipur district and the first split of Raipur district, in 1998-99, possibly as a result of policy paralysis emerging out of shock due to above-said diversification along with a split in the district. Despite these, the production of rice in Amalgamated Raipur in 2014-15 has now crossed the peak of rice production compared to what it was just before the start of diversification, in 1998-99, of rice area to other crops in Raipur district.

ACKNOWLEDGEMENT

The authors acknowledge the staff of Land Records of Chhattisgarh and publications of erstwhile Land Record of Madhya Pradesh, who have provided the secondary data for Rice crop used in the study. Further, this research work has been successfully completed with facilities extended by the Department of Agricultural Statistics and Social Science, Indira Gandhi Krishi Vishwavidyalaya, Raipur.

REFERENCES

1. *Anonymous(2016). Autoregressive integrated moving average, https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average.*
2. *Anonymous (1981 to 1998). The Basic Agricultural, The Commissioner, Land Records and Settlement, Gwalior, Madhya Pradesh.*
3. *Coghlan, A. (2014). A little book of R for Time Series. Release 0.2. Parasite Genomics Group, Wellcome Trust Sanger Institute, Cambridge, U.K.*

4. *Coghlan, A. (2017). A little book of R for Time Series. Release 0.2. Parasite Genomics Group, Wellcome Trust Sanger Institute, Cambridge, U.K.*

5. *H. K. Niranjan et al., Past and Future Trends Using Arima Model in Area, Production and Productivity of Onion in Madhya Pradesh, International Journal of Agricultural Science and Research (IJASR), Volume 6, Issue 2, March - April 2016, pp. 189-194*

6. *Qbal, N. Bakhsh, K., Maqbool, A., and Ahmad, A.S. (2005). Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan. http://www.ijabjass.org.*

7. *Prajakta S. (2004). Time series Forecasting using Holt-Winter Exponential. Kanwal Rekhi School of Information Technology.*

8. *Anonymous (2017). The R Project for Statistical Computing, www.r-project.org.*

9. *Anonymous (2017). Department of Agriculture, Chhattisgarh, www.agridept.cg.gov.in*

APPENDIX-A

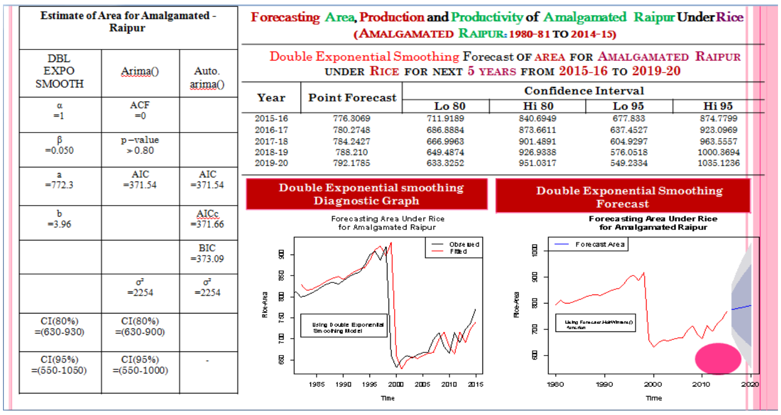


Figure 1: Estimated Forecast Models of Area of Amalgamated Raipur Under Rice for 1980-81 to 2014-15: (a) Observed vs. Fitted and (b) Double Exponential Smoothing Forecast

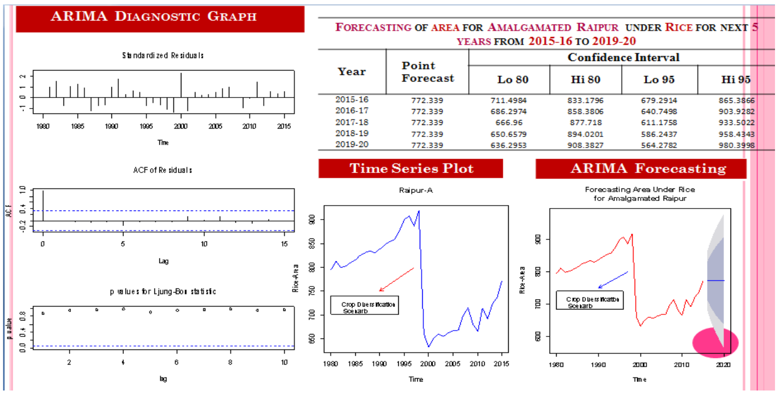
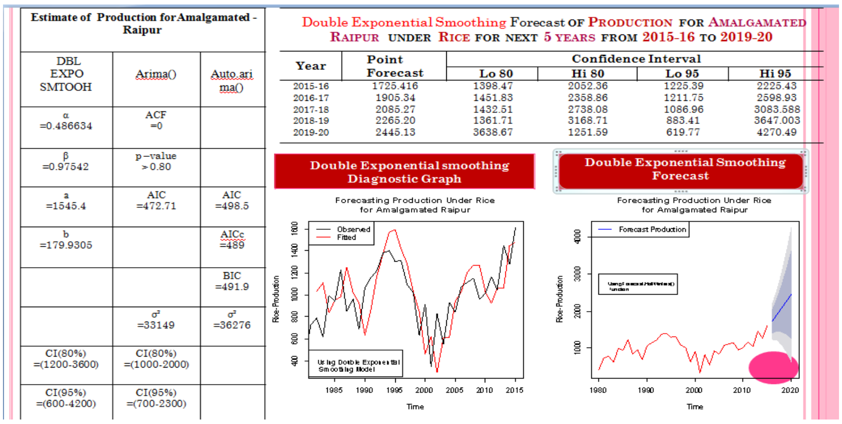


Figure 2: Estimated Forecast Models of Area of Amalgamated Raipur under Rice for 1980-81 to 2014-15: (c) Time Series and (d) ARIMA Forecasting

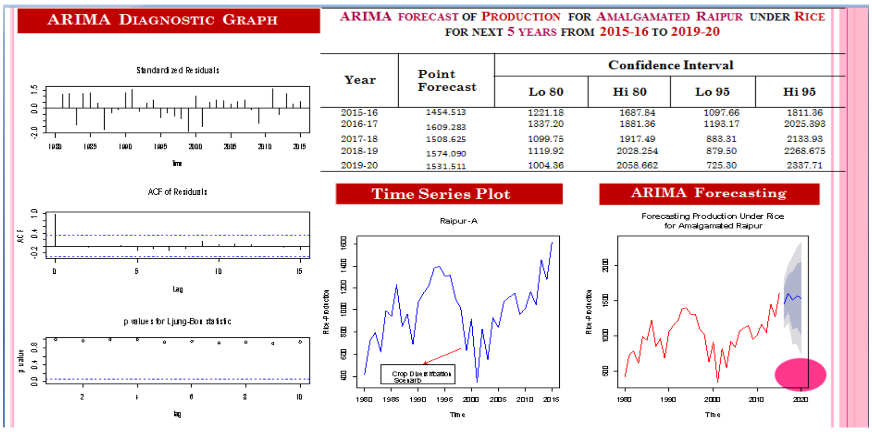


Estimate of Production for Amalgamated - Raipur		
DBL EXPO SMTOOH	Arima()	Auto.arima()
α =0.486634	ACF =0	
β =0.97542	p−value >0.80	
a =1545.4	AIC =472.71	AIC =498.5
b =179.9305		AICc =489
		BIC =491.9
	σ² =33149	σ² =36276
CI(80%) =(1200-3600)	CI(80%) =(1000-2000)	
CI(95%) =(600-4200)	CI(95%) =(700-2300)	

Double Exponential Smoothing Forecast of Production for Amalgamated Raipur under Rice for next 5 years from 2015-16 to 2019-20

Year	Point Forecast	Confidence Interval			
		Lo 80	Hi 80	Lo 95	Hi 95
2015-16	1725.416	1398.47	2052.36	1225.39	2225.43
2016-17	1905.34	1451.83	2358.86	1211.75	2598.93
2017-18	2085.27	1432.51	2738.08	1086.96	3083.588
2018-19	2265.20	1361.71	3168.71	883.41	3647.003
2019-20	2445.13	3638.67	1251.59	619.77	4270.49

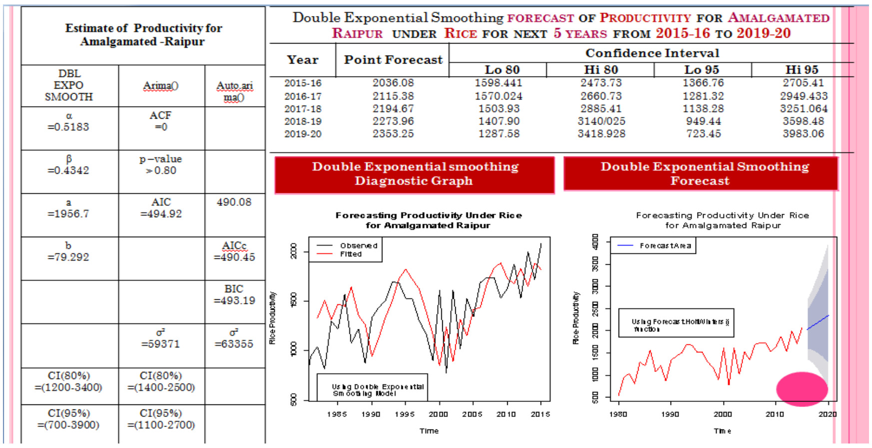
Figure 3: Estimated Forecast Models of Production of Amalgamated Raipur under Rice for 1980-81 to 2014-15: (a) Observed vs. Fitted and (b) Double Exponential Smoothing Forecast



ARIMA forecast of Production for Amalgamated Raipur under Rice for next 5 years from 2015-16 to 2019-20

Year	Point Forecast	Confidence Interval			
		Lo 80	Hi 80	Lo 95	Hi 95
2015-16	1454.513	1221.18	1687.84	1097.66	1811.36
2016-17	1609.283	1337.20	1881.36	1193.17	2025.393
2017-18	1508.625	1099.75	1917.49	883.31	2133.93
2018-19	1574.090	1119.92	2028.254	879.50	2268.675
2019-20	1531.511	1004.36	2058.662	725.30	2337.71

Figure 4: Estimated Forecast Models of Production of Amalgamated Raipur Under Rice for 1980-81 to 2014-15: (c) Time Series and (d) ARIMA Forecasting

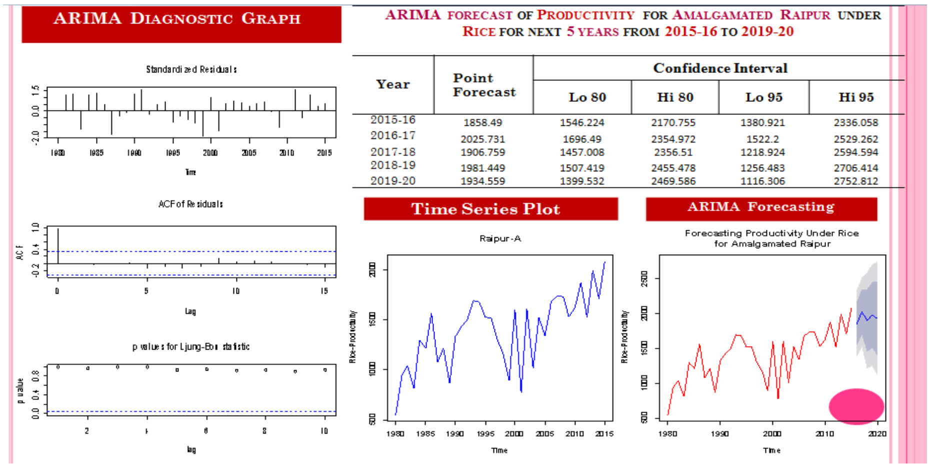


Estimate of Productivity for Amalgamated -Raipur		
DBL EXPO SMOOTH	Arima()	Auto.arima()
α =0.5183	ACF =0	
β =0.4342	p−value >0.80	
a =1956.7	AIC =494.92	490.08
b =79.292		AICc =490.45
		BIC =493.19
	σ² =59371	σ² =63355
CI(80%) =(1200-3400)	CI(80%) =(1400-2500)	
CI(95%) =(700-3900)	CI(95%) =(1100-2700)	

Double Exponential Smoothing forecast of Productivity for Amalgamated Raipur under Rice for next 5 years from 2015-16 to 2019-20

Year	Point Forecast	Confidence Interval			
		Lo 80	Hi 80	Lo 95	Hi 95
2015-16	2036.08	1598.441	2473.73	1366.76	2705.41
2016-17	2115.38	1570.024	2660.73	1281.32	2949.433
2017-18	2194.67	1503.93	2885.41	1138.28	3251.064
2018-19	2273.96	1407.90	3140/025	949.44	3598.48
2019-20	2353.25	1287.58	3418.928	723.45	3983.06

Figure 5: Estimated Forecast Models of Productivity of Amalgamated Raipur Under Rice for 1980-81 to 2014-15: (a) Observed vs. Fitted and (b) Double Exponential Smoothing Forecast



Year	Point Forecast	Confidence Interval			
		Lo 80	Hi 80	Lo 95	Hi 95
2015-16	1858.49	1546.224	2170.755	1380.921	2336.058
2016-17	2025.731	1696.49	2354.972	1522.2	2529.262
2017-18	1906.759	1457.008	2356.51	1218.924	2594.594
2018-19	1981.449	1507.419	2455.478	1256.483	2706.414
2019-20	1934.559	1399.532	2469.586	1116.306	2752.812

Figure 6: Estimated Forecast Models of Productivity of Amalgamated Raipur under Rice for 1980-81 to 2014